#8
Dont
8-7-01

Attorney's Docket No. 5470-130DV

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re: French et al.
Serial No.: 09/497,822
Filed: February 3, 2000
For: *ANDROGEN RECEPTOR PROTEINS,RECOMBINANT DNA MOLECULES CODING FOR SUCH, AND USE OF SUCH COMPOSITIONS*

Examiner: M. Pak
Group Art Unit: 1646

Date: July 31, 2001

Commissioner for Patents
Washington, DC 20231

SUBMITTAL OF FORMAL DRAWINGS

Sir:

Enclosed herewith please find one set (23 sheets) of new formal drawings. It is requested that these new drawings be substituted for the originally filed formal drawings.

Respectfully submitted,

Karen A. Magri
Registration No. 41,965

Customer Number:

20792
PATENT TRADEMARK OFFICE

CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Commissioner For Patents, Washington, DC 20231, on July 31, 2001.

Traci A. Brown
Date of Signature: July 31, 2001

OLIGO A			
COMPLEMENT 5'-		ACC TGT GAG GGC TGT AAG GTC TTC TTC AAA AG -3'	(100%) (**SEQ ID NO:1**)
		** *** * * ** *** *** *** *** *** **	
hAR	(X)	ACA TGT GGA AGC TGC AAG GTC TTC TTC AAA AG	(84%) (**SEQ ID NO:2**)
hPR	(11)	ACC TGT GGG AGC TGT AAG GTC TTC TTT AAG AG	(88%) (**SEQ ID NO:3**)
hMR	(4)	ACC TGT GGC AGC TGC AAA GTT TTC TTC AAA AG	(81%) (**SEQ ID NO:4**)
hGR	(5)	ACT TGT GGA AGC TGT AAA GTT TTC TTC AAA AG	(81%) (**SEQ ID NO:5**)
hER	(6)	TCC TGT GAG GGC TGT AAG GCC TTC TTC AAG AG	(91%) (**SEQ ID NO:6**)
hT3R	(3, 17)	ACG TGT GAA GGC TGC AAG GGT TTC TTT AGA AG	(78%) (**SEQ ID NO:7**)
hRAR	(17)	GCC TGT GAG GGC TGC AAG GGC TTC TTC CGC CG	(78%) (**SEQ ID NO:8**)

FIG. 1A

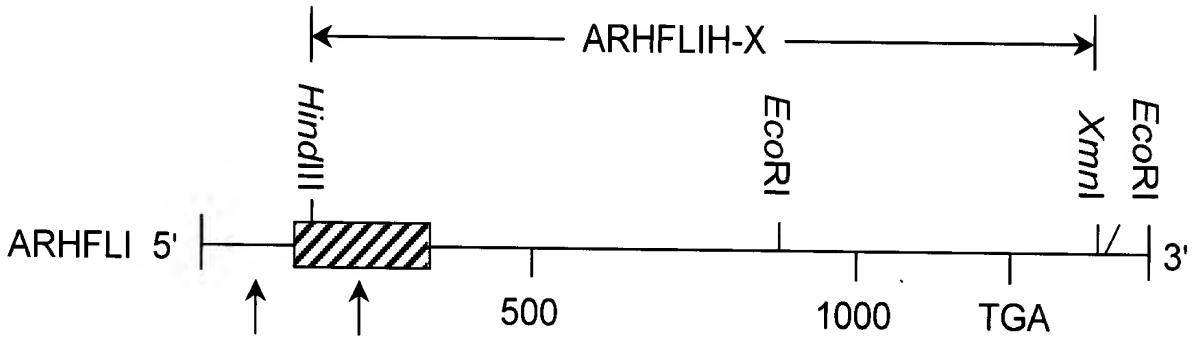


FIG. 1B

DNA-BINDING DOMAIN

			+ + 10 + 20+ 30	
hAR		**(SEQ ID NO:9)**	C L I C G D E A S G C H Y G A L T C G S C K V F F K R A A E G	(100%)
hPR	(AA 567)	**(SEQ ID NO:10)**	C L I C G D E A S G C H Y G V L T C G S C K V F F K R A M E G	(94%)
hMR	(AA 603)	**(SEQ ID NO:11)**	C L V C G D E A S G C H Y G V V T C G S C K V F F K R A V E G	(87%)
hGR	(AA 421)	**(SEQ ID NO:12)**	C L V C S D E A S G C H Y G V L T C G S C K V F F K R A V E G	(87%)
hER	(AA 185)	**(SEQ ID NO:13)**	C A V C N D Y A S G Y H Y G V W S C E G C K A F F K R S I Q G	(55%)
cVDR		**(SEQ ID NO:14)**	C G V C G D R A T G F H F N A M T C E G C K G F F R R S M K R	(48%)
hT3R	(AA 102)	**(SEQ ID NO:15)**	C V V C G D K A T G Y H Y R C I T C E G C K G F F R R T I Q K	(48%)
vERBA	(AA 37)	**(SEQ ID NO:16)**	C V V C G D K A T G Y H Y R C I T C E G C K S F F R R T I Q K	(48%)
hRAR	(AA 58)	**(SEQ ID NO:17)**	C F V C Q D K S S G Y H Y G V S A C E G C K G F F R R S I Q K	(45%)

		+ 40 + 50 + + 60+	
hAR	**(SEQ ID NO:9)**	K Q K Y L C A S R N D C T I D K F R R K N C P S C R L R K C Y E A G M	(100%)
hPR	**(SEQ ID NO:10)**	Q H N Y L C A G R N D C I V D K I R R K N C P A C R L R K C C Q A G M	(71%)
hMR	**(SEQ ID NO:11)**	Q H N Y L C A G R N D C I I D K I R R K N C P A C R L Q K C L Q A G M	(71%)
hGR	**(SEQ ID NO:12)**	Q H N Y L C A G R N D C I I D K I R R K N C P A C R Y R K C L Q A G M	(71%)
hER	**(SEQ ID NO:13)**	H N D Y M C P A T N Q C T I D K N R R K S C Q A C R L R K C Y E V G M	(63%)
cVDR	**(SEQ ID NO:14)**	K A M F T C P F N G D C K I T K D N R R H C Q A C R L K R C V D I G M	(40%)
hT3R	**(SEQ ID NO:15)**	N L H P S Y S C K Y E G K C V I D K V T R N Q C Q E C R F K K C I Y V G M	(40%)
vERBA	**(SEQ ID NO:16)**	N L H P T T S C T Y D G C C V I D K I T R N Q C Q L C R F K K C I S V G M	(37%)
hRAR	**(SEQ ID NO:17)**	N M V Y T C H R D K N C I I N K V T R N R C Q Y C R L Q K C F E V G M	(43%)

FIG. 1C

FIG. 2A

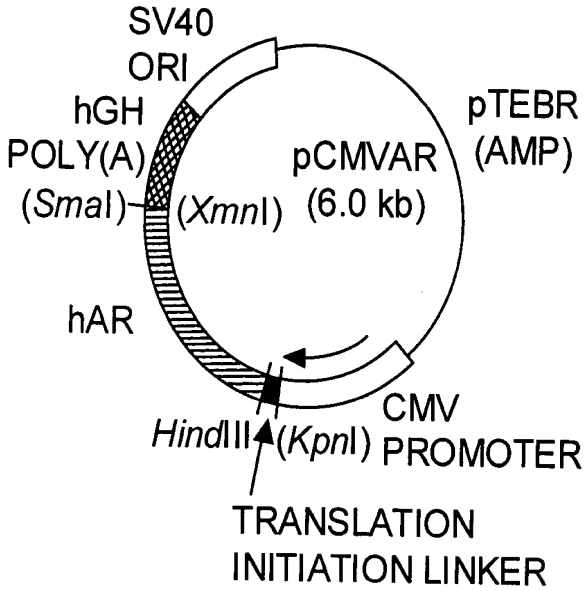


FIG. 2B

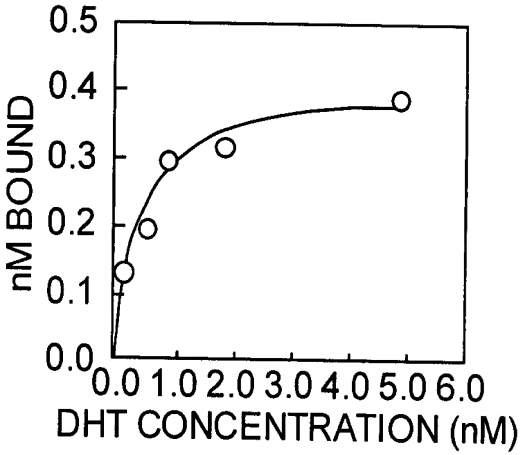


FIG. 2C

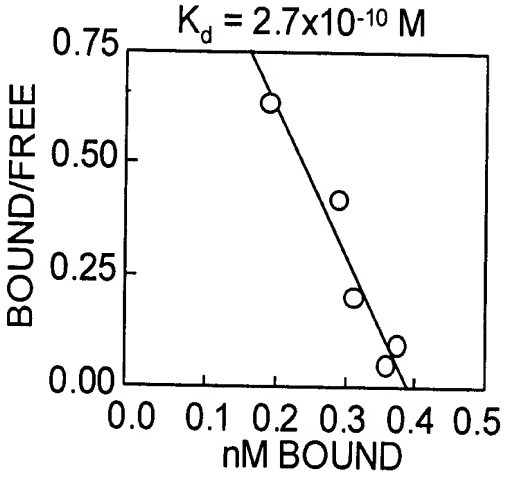
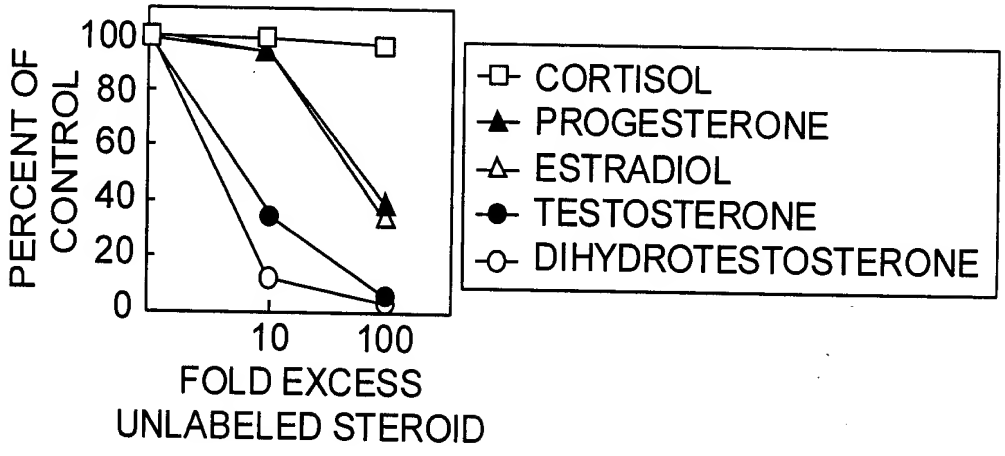


FIG. 2D



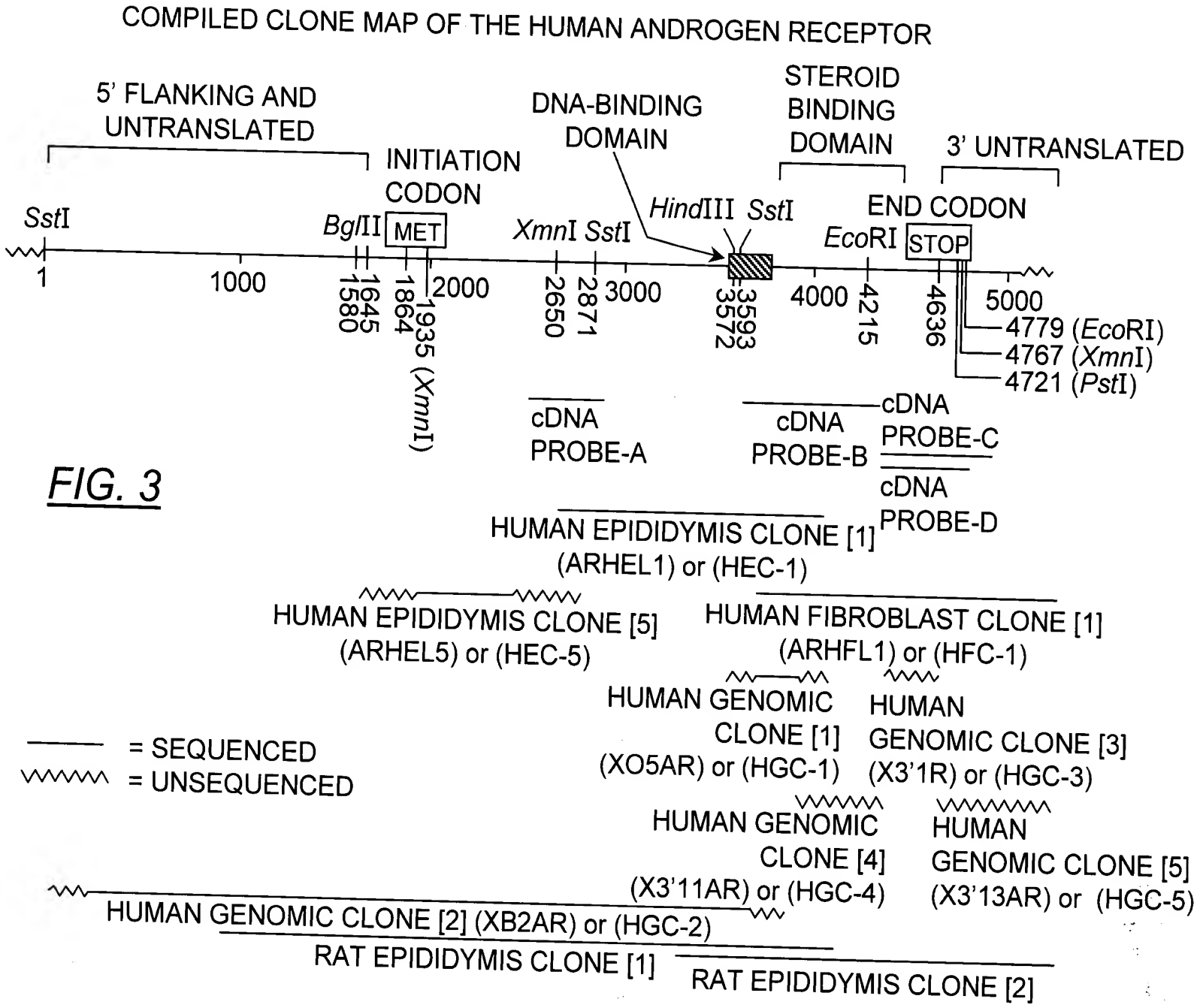

COMPILED CLONE MAP OF THE HUMAN ANDROGEN RECEPTOR
5' FLANKING AND UNTRANSLATED
INITIATION CODON
DNA-BINDING DOMAIN
STEROID BINDING DOMAIN
3' UNTRANSLATED
END CODON
SstI
BglII
MET
XmnI
SstI
HindIII
SstI
EcoRI
STOP
1
1000
1580
1645
1864
2000
1935 (XmnI)
2650
2871
3000
3572
3593
4000
4215
4636
5000
4779 (EcoRI)
4767 (XmnI)
4721 (PstI)
cDNA PROBE-A
cDNA PROBE-B
cDNA PROBE-C
cDNA PROBE-D
HUMAN EPIDIDYMIS CLONE [1] (ARHEL1) or (HEC-1)
HUMAN EPIDIDYMIS CLONE [5] (ARHEL5) or (HEC-5)
HUMAN FIBROBLAST CLONE [1] (ARHFL1) or (HFC-1)
HUMAN GENOMIC CLONE [1] (XO5AR) or (HGC-1)
HUMAN GENOMIC CLONE [3] (X3'1R) or (HGC-3)
HUMAN GENOMIC CLONE [4] (X3'11AR) or (HGC-4)
HUMAN GENOMIC CLONE [5] (X3'13AR) or (HGC-5)
HUMAN GENOMIC CLONE [2] (XB2AR) or (HGC-2)
RAT EPIDIDYMIS CLONE [1]
RAT EPIDIDYMIS CLONE [2]
= SEQUENCED
= UNSEQUENCED

FIG. 3

```
          10                  30                  50
          .                   .                   .
GAGCTCTGGACAAAATTGAGCGCCTATGTGTACATGGCAAGTGTTTTTAGTGTTTGTGTG
CTCGAGACCTGTTTTAACTCGCGGATACACATGTACCGTTCACAAAAATCACAAACACAC

          70                  90                  110
          .                   .                   .
TTTACCTGCTTGTCTGGGTGATTTTGCCTTTGAGAGTCTGGATGAGAAATGCATGGTTAA
AAATGGACGAACAGACCCACTAAAACGGAAACTCTCAGACCTACTCTTTACGTACCAATT

          130                 150                 170
          .                   .                   .
AGGCAATTCCAGACAGGAAGAAAGGCAGAGAAGAGGGTAGAAATGACCTCTGATTCTTGG
TCCGTTAAGGTCTGTCCTTCTTTCCGTCTCTTCTCCCATCTTTACTGGAGACTAAGAACC

          190                 210                 230
          .                   .                   .
GGCTGAGGGTTCCTAGAGCAAATGGCACAATGCCACGAGGCCCGATCTATCCCTATGACG
CCGACTCCCAAGGATCTCGTTTACCGTGTTACGGTGCTCCGGGCTAGATAGGGATACTGC

          250                 270                 290
          .                   .                   .
GAACTCTAAGGTTTCAGCATCAGCTATCTGCTGGCTTGGTCACTGGCTTGCCTCCTCAGT
CTTGAGATTCCAAAGTCGTAGTCGATAGACGACCGAACCAGTGACCGAACGGAGGAGTCA

          310                 330                 350
          .                   .                   .
TTGTAGGAGACTCTCCCACTCTCCCATCTGCGCGCTCTTATCAGTCCTGAAAAGAACCCN
AACATCCTCTGAGAGGGTGAGAGGGTAGACGCGCGAGAATAGTCAGGACTTTTCTTGGGN

          370                 390                 410
          .                   .                   .
TGGCNAGCCAGGAGCNAGGTATTCNTATCGTCCTTTTCNTCCTCCTNGCCTCACCTNGTT
ACCGNTCGGTCCTCGNTCCATAAGNATAGCAGGAAAAGNAGGAGGANCGGAGTGGANCAA

          430                 450                 470
          .                   .                   .
GNTTTTTAGATTGGNCTTNGNAACCAAATTGTATGCTGGCCTCCAGGAAATCTGGAGCC
CNAAAAATCTAACCNGAANCNTTGGTTTAAACATACGACCGGAGGTCCTTTAGACCTCGG

          490                 510                 530
          .                   .                   .
TGGCGCCTAAACCTTGGTTTAGGAAAGCAGGAGCTATTCAGGAAGCAGGGTCCTCCAGGG
ACCGCGGATTTGGAACCAAATCCTTTCGTCCTCGATAAGTCCTTCGTCCCAGGAGGTCCC

          550                 570                 590
          .                   .                   .
CTAGAGCTAGCCTCTCCTGCCCTCGCCCACGTGCGCCAGCACTTGTTTCTCCAAAGCNAC
GATCTCGATCGGAGAGGACGGGAGCGGGTGCACGCGGTCGTGAACAAAGAGGTTTCGNTG
```

FIG. 4A

610 630 650

TAGGCAGGCGTTAGCGCGCGGTGAGGGGAGGGGAGAAAAGGAAAGGGGAGGGGAGGGAAA
ATCCGTCCGCAATCGCGCGCCACTCCCCTCCCCTCTTTTCCTTTCCCCTCCCCTCCCTTT

670 690 710

AGGAGGTGGGAAGGCAAGGAGGCCGGCCNGGTGGGGGCGGGACCCGACTCGCANNAACTG
TCCTCCACCCTTCCGTTCCTCCGGCCGGNCCACCCCCGCCCTGGGCTGAGCGTNNTTGAC

730 750 770

TTGCATTTGCTCTCCACCTCCCAGCGCCCCCTCCGAGATCCCGGGGAGCCAGCTTGCTGG
AACGTAAACGAGAGGTGGAGGGTCGCGGGGGAGGCTCTAGGGCCCCTCGGTCGAACGACC

790 810 830

GAGAGCGGGAACGGTCCGGAGCAAGCCCAGAGGCAGAGGAGGCGACAGAGGGAAAAAGGG
CTCTCGCCCTTGCCAGGCCTCGTTCGGGTCTCCGTCTCCTCCGCTGTCTCCCTTTTTCCC

850 870 890

CCCNAGCTAGCCGCTCCAGTGCTGTACAGNAGCCGAAGGACGCACCACGCCAGCCCCAGC
GGGNTCGATCGGCGAGGTCACGACATGTCNTCGGCTTCCTGCGTGGTGCGGTCGGGGTCG

910 930 950

CCGGCTCCAGCGACAGCNAACGCCTCTTGCANGCGTTCGAAGCCGCCGCCCGGAGCTGCC
GGCCGAGGTCGCTGTCGNTTGCGGAGAACGTNCGCAAGCTTCGGCGGCGGGCCTCGACGG

970 990 1010

CTTTCCTCTTCGGTGAAGTTTTTAAAAGCTGCTAAAGACTCGGAGGAAGCAAGGAAAGTG
GAAAGGAGAAGCCACTTCAAAAATTTTCGACGATTTCTGAGCCTCCTTCGTTCCTTTCAC

1030 1050 1070

CCTGGTAGGACTGACGGCTGCCTTTGTCCTCCTCCTCTCCACCCCGCCTCCCCCCACCCT
GGACCATCCTGACTGCCGACGGAAACAGGAGGAGGAGAGGTGGGGCGGAGGGGGGTGGGA

1090 1110 1130

GCCTTCCCCCCCTCCCCCGTCTTCTCTCCCGCAGCTGCCTCAGTCGGCTACTCTCAGCCA
CGGAAGGGGGGGAGGGGGCAGAAGAGAGGGCGTCGACGGAGTCAGCCGATGAGAGTCGGT

1150 1170 1190

ACCCCCCTCACCACCCTTCTCCCCACCCGCCCCCCCGCCCCCGTCGGCCCAGCGNTGNCA
TGGGGGGAGTGGTGGGAAGAGGGGTGGGCGGGGGGGCGGGGGCAGCCGGGTCGCNACNGT

FIG. 4B

1210 1230 1250

GNCCGAGTTTGCAGAGAGGTAACTCCCTTTGGCTGCGAGCGGGCGAGNCTAGCTGCACAT
CNGGCTCAAACGTCTCTCCATTGAGGGAAACCGACGCTCGCCCGCTCNGATCGACGTGTA

1270 1290 1310

TGCAAAGAAGGCTCTTAGGAGCAGGCGACTGGGGAGCGGCTTCAGCACTGCAGCCACGAC
ACGTTTCTTCCGAGAATCCTCGTCCGCTGACCCCTCGCCGAAGTCGTGACGTCGGTGCTG

1330 1350 1370

CNGCCTGGTTAGGCTGCACGCGGAGAGAACCCTCTGTTTTCCCCCACTCTCTCTCCACCT
GNCGGACCAATCCGACGTGCGCCTCTCTTGGGAGACAAAAGGGGGTGAGAGAGAGGTGGA

1390 1410 1430

CCTCCTGCCTTCCCCACCCCGAGTGCGGAGCCAGAGATCAAAAGATGAAAAGGCAGTCAG
GGAGGACGGAAGGGGTGGGGCTCACGCCTCGGTCTCTAGTTTTCTACTTTTCCGTCAGTC

1450 1470 1490

GTCTTCAGTAGCCAAAAAACAAAACAAACAAAAACAAAAAAGCCGAAATAAAAGAAAAAG
CAGAAGTCATCGGTTTTTTGTTTTGTTTGTTTTTGTTTTTTCGGCTTTATTTTCTTTTTC

1510 1530 1550

ATAATAACTCAGTTCTTATTTGCACCTACTTCAGTGGACACTGAATTTGGAAGGTGGAGG
TATTATTGAGTCAAGAATAAACGTGGATGAAGTCACCTGTGACTTAAACCTTCCACCTCC

1570 1590 1610

ATTTTGTTTTTTTCTTTTAAGATCTGGGCATCTTTTGAATCTACCCTTCAAGTATTAAGA
TAAAACAAAAAAAGAAAATTCTAGACCCGTAGAAAACTTAGATGGGAAGTTCATAATTCT

1630 1650 1670

GACAGACTGTGAGCCTAGCAGGGCAGATCTTGTCCACCGTGTGTCTTCTTCTGCACGAGA
CTGTCTGACACTCGGATCGTCCCGTCTAGAACAGGTGGCACACAGAAGAAGACGTGCTCT

1690 1710 1730

CTTTGAGGCTGTCAGAGCGCTTTTTGCGTGGTTGCTCCCGCAAGTTTCCTTCTCTGGAGC
GAAACTCCGACAGTCTCGCGAAAAACGCACCAACGAGGGCGTTCAAAGGAAGAGACCTCG

1750 1770 1790

TTCCCGCAGGTGGGCAGCTAGCTGCAGCGACTACCGCATCATCACAGCCTGTTGAACTCT
AAGGGCGTCCACCCGTCGATCGACGTCGCTGATGGCGTAGTAGTGTCGGACAACTTGAGA

FIG. 4C

1810 1830 1850

TCTGAGCAAGAGAAGGGGAGGCGGGGTAAGGGAAGTAGGTGGAAGATTCAGCCAAGCTCA
AGACTCGTTCTCTTCCCCTCCGCCCCATTCCCTTCATCCACCTTCTAAGTCGGTTCGAGT

1870 1890 1910

AGGATGGAAGTGCAGTTAGGGCTGGGAAGGGTCTACCCTCGGCCGCCGTCCAAGACCTAC
TCCTACCTTCACGTCAATCCCGACCCTTCCCAGATGGGAGCCGGCGGCAGGTTCTGGATG

1930 1950 1970

CGAGGAGCTTTCCAGAATCTGTTCCAGAGCGTGCGCGAAGTGATCCAGAACCCGGGCCCC
GCTCCTCGAAAGGTCTTAGACAAGGTCTCGCACGCGCTTCACTAGGTCTTGGGCCCGGGG

1990 2010 2030

AGGCACCCAGAGGCCGCGAGCGCAGCACCTCCCGGCGCCAGTTTGCTGCTGCTGCAGCAG
TCCGTGGGTCTCCGGCGCTCGCGTCGTGGAGGGCCGCGGTCAAACGACGACGACGTCGTC

2050 2070 2090

CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG
GTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC

2110 2130 2150

CAGCAGCAAGAGACTAGCCCCAGGCAGCAGCAGCAGCAGGGTGAGGATGGTTCTCCC
GTCGTCGTTCTCTGATCGGGGTCCGTCGTCGTCGTCGTCCCACTCCTACCAAGAGGG

2170 2190 2210

CAAGCCCATCGTAGAGGCCCCACAGGCTACCTGGTCCTGGATGAGGAACAGCAACCTTCA
GTTCGGGTAGCATCTCCGGGGTGTCCGATGGACCAGGACCTACTCCTTGTCGTTGGAAGT

2230 2250 2270

CAGCCGCAGTCGGCCCTGGAGTGCCACCCCGAGAGAGGTTGCGTCCCAGAGCCTGGAGCC
GTCGGCGTCAGCCGGGACCTCACGGTGGGGCTCTCTCCAACGCAGGGTCTCGGACCTCGG

2290 2310 2330

GCCGTGGCCGCCAGCAAGGGGCTGCCGCAGCAGCTGCCAGCACCTCCGGACGAGGATGAC
CGGCACCGGCGGTCGTTCCCCGACGGCGTCGTCGACGGTCGTGGAGGCCTGCTCCTACTG

2350 2370 2390

TCAGCTGCCCCATCCACGTTGTCCCTGCTGGGCCCCACTTTCCCCGGCTTAAGCAGCTGC
AGTCGACGGGGTAGGTGCAACAGGGACGACCCGGGGTGAAAGGGGCCGAATTCGTCGACG

FIG. 4D

2410 2430 2450

TCCGCTGACCTTAAAGACATCCTGAGCGAGGCCAGCACCATGCAACTCCTTCAGCAACAG
AGGCGACTGGAATTTCTGTAGGACTCGCTCCGGTCGTGGTACGTTGAGGAAGTCGTTGTC

2470 2490 2510

CAGCAGGAAGCAGTATCCGAAGGCAGCAGCAGCGGGAGAGCGAGGGAGGCCTCGGGGGCT
GTCGTCCTTCGTCATAGGCTTCCGTCGTCGTCGCCCTCTCGCTCCCTCCGGAGCCCCCGA

2530 2550 2570

CCCACTTCCTCCAAGGACAATTACTTAGGGGGCACTTCGACCATTTCTGACAACGCCAAG
GGGTGAAGGAGGTTCCTGTTAATGAATCCCCCGTGAAGCTGGTAAAGACTGTTGCGGTTC

2590 2610 2630

GAGTTGTGTAAGGCAGTGTCGGTGTCCATGGGCCTGGGTGTGGAGGCGTTGGAGCATCTG
CTCAACACATTCCGTCACAGCCACAGGTACCCGGACCCACACCTCCGCAACCTCGTAGAC

2650 2670 2690

AGTCCAGGGGAACAGCTTCGGGGGGATTGCATGTACGCCCCACTTTTGGGAGTTCCACCC
TCAGGTCCCCTTGTCGAAGCCCCCCTAACGTACATGCGGGGTGAAAACCCTCAAGGTGGG

2710 2730 2750

GCTGTGCGTCCCACTCCTTGTGCCCCATTGGCCGAATGCAAAGGTTCTCTGCTAGACGAC
CGACACGCAGGGTGAGGAACACGGGGTAACCGGCTTACGTTTCCAAGAGACGATCTGCTG

2770 2790 2810

AGCGCAGGCAAGAGCACTGAAGATACTGCTGAGTATTCCCCTTTCAAGGGAGGTTACACC
TCGCGTCCGTTCTCGTGACTTCTATGACGACTCATAAGGGGAAAGTTCCCTCCAATGTGG

2830 2850 2870

AAAGGGCTAGAAGGCGAGAGCCTAGGCTGCTCTGGCAGCGCTGCAGCAGGGAGCTCCGGG
TTTCCCGATCTTCCGCTCTCGGATCCGACGAGACCGTCGCGACGTCGTCCCTCGAGGCCC

2890 2910 2930

ACACTTGAACTGCCGTCTACCCTGTCTCTCTACAAGTCCGGAGCACTGGACGAGGCAGCT
TGTGAACTTGACGGCAGATGGGACAGAGAGATGTTCAGGCCTCGTGACCTGCTCCGTCGA

2950 2970 2990

GCGTACCAGAGTCGCGACTACTACAACTTTCCACTGGCTCTGGCCGGACCGCCGCCCCCT
CGCATGGTCTCAGCGCTGATGATGTTGAAAGGTGACCGAGACCGGCCTGGCGGCGGGGGA

FIG. 4E

3010 3030 3050

CCGCCGCCTCCCCATCCCCACGCTCGCATCAAGCTGGAGAACCCGCTGGACTACGGCAGC
GGCGGCGGAGGGGTAGGGGTGCGAGCGTAGTTCGACCTCTTGGGCGACCTGATGCCGTCG

3070 3090 3110

GCCTGGGCGGCTGCGGCGGCGCAGTGCCGCTATGGGGACCTGGCGAGCCTGCATGGCGCG
CGGACCCGCCGACGCCGCCGCGTCACGGCGATACCCCTGGACCGCTCGGACGTACCGCGC

3130 3150 3170

GGTGCAGCGGGACCCGGTTCTGGGTCACCCTCAGCCGCCGCTTCCTCATCCTGGCACACT
CCACGTCGCCCTGGGCCAAGACCCAGTGGGAGTCGGCGGCGAAGGAGTAGGACCGTGTGA

3190 3210 3230

CTCTTCACAGCCGAAGAAGGCCAGTTGTATGGACCGTGTGGTGGTGGTGGGGGTGGTGGC
GAGAAGTGTCGGCTTCTTCCGGTCAACATACCTGGCACACCACCACCACCCCCACCACCG

3250 3270 3290

GGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGAGGCGGGA
CCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCTCCGCCCT

3310 3330 3350

GCTGTAGCCCCCTACGGCTACACTCGGCCCCCTCAGGGGCTGGCGGGCCAGGAAAGCGAC
CGACATCGGGGGATGCCGATGTGAGCCGGGGGAGTCCCCGACCGCCCGGTCCTTTCGCTG

3370 3390 3410

TTCACCGCACCTGATGTGTGGTACCCTGGCGGCATGGTGAGCAGAGTGCCCTATCCCAGT
AAGTGGCGTGGACTACACACCATGGGACCGCCGTACCACTCGTCTCACGGGATAGGGTCA

3430 3450 3470

CCCACTTGTGTCAAAAGCGAAATGGGCCCCTGGATGGATAGCTACTCCGGACCTTACGGG
GGGTGAACACAGTTTTCGCTTTACCCGGGGACCTACCTATCGATGAGGCCTGGAATGCCC

3490 3510 3530

GACATGCGTTTGGAGACTGCCAGGGACCATGTTTTGCCCATTGACTATTACTTTCCACCC
CTGTACGCAAACCTCTGACGGTCCCTGGTACAAAACGGGTAACTGATAATGAAAGGTGGG

3550 3570 3590

CAGAAGACCTGCCTGATCTGTGGAGATGAAGCTTCTGGGTGTCACTATGGAGCTCTCACA
GTCTTCTGGACGGACTAGACACCTCTACTTCGAAGACCCACAGTGATACCTCGAGAGTGT

FIG. 4F

3610 3630 3650

TGTGGAAGCTGCAAGGTCTTCTTCAAAAGAGCCGCTGAAGGGAAACAGAAGTACCTGTGC
ACACCTTCGACGTTCCAGAAGAAGTTTTCTCGGCGACTTCCCTTTGTCTTCATGGACACG

3670 3690 3710

GCCAGCAGAAATGATTGCACTATTGATAAATTCCGAAGGAAAAATTGTCCATCTTGTCGT
CGGTCGTCTTTACTAACGTGATAACTATTTAAGGCTTCCTTTTTAACAGGTAGAACAGCA

3730 3750 3770

CTTCGGAAATGTTATGAAGCAGGGATGACTCTGGGAGCCCGGAAGCTGAAGAAACTTGGT
GAAGCCTTTACAATACTTCGTCCCTACTGAGACCCTCGGGCCTTCGACTTCTTTGAACCA

3790 3810 3830

AATCTGAAACTACAGGAGGAAGGAGAGGCTTCCAGCACCACCAGCCCCACTGAGGAGACA
TTAGACTTTGATGTCCTCCTTCCTCTCCGAAGGTCGTGGTGGTCGGGGTGACTCCTCTGT

3850 3870 3890

ACCCAGAAGCTGACAGTGTCACACATTGAAGGCTATGAATGTCAGCCCATCTTTCTGAAT
TGGGTCTTCGACTGTCACAGTGTGTAACTTCCGATACTTACAGTCGGGTAGAAAGACTTA

3910 3930 3950

GTCCTGGAAGCCATTGAGCCAGGTGTAGTGTGTGCTGGACACGACAACAACCAGCCCGAC
CAGGACCTTCGGTAACTCGGTCCACATCACACACGACCTGTGCTGTTGTTGGTCGGGCTG

3970 3990 4010

TCCTTTGCAGCCTTGCTCTCTAGCCTCAATGAACTGGGAGAGAGACAGCTTGTACACGTG
AGGAAACGTCGGAACGAGAGATCGGAGTTACTTGACCCTCTCTCTGTCGAACATGTGCAC

4030 4050 4070

GTCAAGTGGGCCAAGGCCTTGCCTGGCTTCCGCAACTTACACGTGGACGACCAGATGGCT
CAGTTCACCCGGTTCCGGAACGGACCGAAGGCGTTGAATGTGCACCTGCTGGTCTACCGA

4090 4110 4130

GTCATTCAGTACTCCTGGATGGGGCTCATGGTGTTTGCCATGGGCTGGCGATCCTTCACC
CAGTAAGTCATGAGGACCTACCCCGAGTACCACAAACGGTACCCGACCGCTAGGAAGTGG

4150 4170 4190

AATGTCAACTCCAGGATGCTCTACTTCGCCCCTGATCTGGTTTTCAATGAGTACCGCATG
TTACAGTTGAGGTCCTACGAGATGAAGCGGGGACTAGACCAAAAGTTACTCATGGCGTAC

FIG. 4G

4210 4230 4250

CACAAGTCCCGGATGTACAGCCAGTGTGTCCGAATGAGGCACCTCTCTCAAGAGTTTGGA
GTGTTCAGGGCCTACATGTCGGTCACACAGGCTTACTCCGTGGAGAGAGTTCTCAAACCT

4270 4290 4310

TGGCTCCAAATCACCCCCCAGGAATTCCTGTGCATGAAAGCACTGCTACTCTTCAGCATT
ACCGAGGTTTAGTGGGGGGTCCTTAAGGACACGTACTTTCGTGACGATGAGAAGTCGTAA

4330 4350 4370

ATTCCAGTGGATGGGCTGAAAAATCAAAAATTCTTTGATGAACTTCGAATGAACTACATC
TAAGGTCACCTACCCGACTTTTTAGTTTTTAAGAAACTACTTGAAGCTTACTTGATGTAG

4390 4410 4430

AAGGAACTCGATCGTATCATTGCATGCAAAAGAAAAAATCCCACATCCTGCTCAAGACGC
TTCCTTGAGCTAGCATAGTAACGTACGTTTTCTTTTTTAGGGTGTAGGACGAGTTCTGCG

4450 4470 4490

TTCTACCAGCTCACCAAGCTCCTGGACTCCGTGCAGCCTATTGCGAGAGAGCTGCATCAG
AAGATGGTCGAGTGGTTCGAGGACCTGAGGCACGTCGGATAACGCTCTCTCGACGTAGTC

4510 4530 4550

TTCACTTTTGACCTGCTAATCAAGTCACACATGGTGAGCGTGGACTTTCCGGAAATGATG
AAGTGAAAACTGGACGATTAGTTCAGTGTGTACCACTCGCACCTGAAAGGCCTTTACTAC

4570 4590 4610

GCAGAGATCATCTCTGTGCAAGTGCCCAAGATCCTTTCTGGGAAAGTCAAGCCCATCTAT
CGTCTCTAGTAGAGACACGTTCACGGGTTCTAGGAAAGACCCTTTCAGTTCGGGTAGATA

4630 4650 4670

TTCCACACCCAGTGAAGCATTGGAAACCCTATTTCCCCACCCCAGCTCATGCCCCCTTTC
AAGGTGTGGGTCACTTCGTAACCTTTGGGATAAAGGGGTGGGGTCGAGTACGGGGGAAAG

4690 4710 4730

AGATGTCTTCTGCCTGTTATAACTCTGCACTACTCCTCTGCAGTGCCTTGGGGAATTTCC
TCTACAGAAGACGGACAATATTGAGACGTGATGAGGAGACGTCACGGAACCCCTTAAAGG

4750 4770 4790

TCTATTGATGTACAGTCTGTCATGAACATGTTCCTGAATTCTATTTGCTGGGCTTTTTTT
AGATAACTACATGTCAGACAGTACTTGTACAAGGACTTAAGATAAACGACCCGAAAAAAA

FIG. 4H

```
        4810                4830                4850
.           .           .           .           .           .
TTCTCTTTCTCTCCTTTCTTTTTCTTCTTCCCTCCCTATCTAACCCTCCCATGGCACCTT
AAGAGAAAGAGAGGAAAGAAAAAGAAGAAGGGAGGGATAGATTGGGAGGGTACCGTGGAA

        4870                4890                4910
.           .           .           .           .           .
CAGACTTTGCTTCCCATTGTGGCTCCTATCTGTGTTTTGAATGGTGTTGTATGCCTTTAA
GTCTGAAACGAAGGGTAACACCGAGGATAGACACAAAACTTACCACAACATACGGAAATT

        4930                4950                4970
.           .           .           .           .           .
ATCTGTGATGATCCTCATATGGCCCAGTGTCAAGTTGTGCTTGTTTACAGCACTACTCTG
TAGACACTACTAGGAGTATACCGGGTCACAGTTCAACACGAACAAATGTCGTGATGAGAC

        4990                5010                5030
.           .           .           .           .           .
TGCCAGCCACACAAACGTTTACTTATCTTATGCCACGGGAAGTTTAGAGAGCTAAGATTA
ACGGTCGGTGTGTTTGCAAATGAATAGAATACGGTGCCCTTCAAATCTCTCGATTCTAAT

        5050                5070
.           .           .           .           .
TCTGGGGAAATCAAAACAAAAAACAAGCAAACAAAAAAAAAA
AGACCCCTTTAGTTTTGTTTTTTGTTCGTTTGTTTTTTTTTT
```

FIG. 4I

```
   1 GAGCTCTGGACAAAATTGAGCGCCTATGTGTACATGGCAAGTGTTTTTAGTGTTTGTGTG
  61 TTTACCTGCTTGTCTGGGTGATTTTGCCTTTGAGAGTCTGGATGAGAAATGCATGGTTAA
 121 AGGCAATTCCAGACAGGAAGAAAGGCAGAGAAGAGGGTAGAAATGACCTCTGATTCTTGG
 181 GGCTGAGGGTTCCTAGAGCAAATGGCACAATGCCACGAGGCCCGATCTATCCCTATGACG
 241 GAACTCTAAGGTTTCAGCATCAGCTATCTGCTGGCTTGGTCACTGGCTTGCCTCCTCAGT
 301 TTGTAGGAGACTCTCCCACTCTCCCATCTGCGCGCTCTTATCAGTCCTGAAAAGAACCCN
 361 TGGCNAGCCAGGAGCNAGGTATTCNTATCGTCCTTTTCNTCCTCCTNGCCTCACCTNGTT
 421 GNTTTTTAGATTGGNCTTNGNAACCAAATTTGTATGCTGGCCTCCAGGAAATCTGGAGCC
 481 TGGCGCCTAAACCTTGGTTTAGGAAAGCAGGAGCTATTCAGGAAGCAGGGTCCTCCAGGG
 541 CTAGAGCTAGCCTCTCCTGCCCTCGCCCACGTGCGCCAGCACTTGTTTCTCCAAAGCNAC
 601 TAGGCAGGCGTTAGCGCGCGGTGAGGGGAGGGGAGAAAAGGAAAGGGGAGGGGAGGGAAA
 661 AGGAGGTGGGAAGGCAAGGAGGCCGGCCNGGTGGGGGCGGGACCCGACTCGCANNAACTG
 721 TTGCATTTGCTCTCCACCTCCCAGCGCCCCCTCCGAGATCCCGGGGAGCCAGCTTGCTGG
 781 GAGAGCGGGAACGGTCCGGAGCAAGCCCAGAGGCAGAGGAGGCGACAGAGGGAAAAAGGG
 841 CCCNAGCTAGCCGCTCCAGTGCTGTACAGNAGCCGAAGGACGCACCACGCCAGCCCCAGC
 901 CCGGCTCCAGCGACAGCNAACGCCTCTTGCANGCGTTCGAAGCCGCCGCCCGGAGCTGCC
 961 CTTTCCTCTTCGGTGAAGTTTTTAAAAGCTGCTAAAGACTCGGAGGAAGCAAGGAAAGTG
1021 CCTGGTAGGACTGACGGCTGCCTTTGTCCTCCTCCTCTCCACCCCGCCTCCCCCCACCCT
1081 GCCTTCCCCCCCTCCCCCGTCTTCTCTCCCGCAGCTGCCTCAGTCGGCTACTCTCAGCCA
1141 ACCCCCCTCACCACCCTTCTCCCCACCCGCCCCCCCGCCCCCGTCGGCCCAGCGNTGNCA
1201 GNCCGAGTTTGCAGAGAGGTAACTCCCTTTGGCTGCGAGCGGGCGAGNCTAGCTGCACAT
1261 TGCAAAGAAGGCTCTTAGGAGCAGGCGACTGGGGAGCGGCTTCAGCACTGCAGCCACGAC
1321 CNGCCTGGTTAGGCTGCACGCGGAGAGAACCCTCTGTTTTCCCCCACTCTCTCTCCACCT
1381 CCTCCTGCCTTCCCCACCCCGAGTGCGGAGCCAGAGATCAAAAGATGAAAAGGCAGTCAG
1441 GTCTTCAGTAGCCAAAAAACAAAACAAACAAAAACAAAAAAGCCGAAATAAAAGAAAAAG
```

FIG. 5A

```
1501 ATAATAACTCAGTTCTTATTTGCACCTACTTCAGTGGACACTGAATTTGGAAGGTGGAGG
1561 ATTTTGTTTTTTTCTTTTAAGATCTGGGCATCTTTTGAATCTACCCTTCAAGTATTAAGA
1621 GACAGACTGTGAGCCTAGCAGGGCAGATCTTGTCCACCGTGTGTCTTCTTCTGCACGAGA
1681 CTTTGAGGCTGTCAGAGCGCTTTTTGCGTGGTTGCTCCCGCAAGTTTCCTTCTCTGGAGC
1741 TTCCCGCAGGTGGGCAGCTAGCTGCAGCGACTACCGCATCATCACAGCCTGTTGAACTCT
1801 TCTGAGCAAGAGAAGGGGAGGCGGGGTAAGGGAAGTAGGTGGAAGATTCAGCCAAGCTCA
1861 AGGATGGAAGTGCAGTTAGGGCTGGGAAGGGTCTACCCTCGGCCGCCGTCCAAGACCTAC
        MetGluValGlnLeuGlyLeuGlyArgValTyrProArgProProSerLysThrTyr
1921 CGAGGAGCTTTCCAGAATCTGTTCCAGAGCGTGCGCGAAGTGATCCAGAACCCGGGCCCC
     ArgGlyAlaPheGlnAsnLeuPheGlnSerValArgGluValIleGlnAsnProGlyPro
1981 AGGCACCCAGAGGCCGCGAGCGCAGCACCTCCCGGCGCCAGTTTGCTGCTGCTGCAGCAG
     ArgHisProGluAlaAlaSerAlaAlaProProGlyAlaSerLeuLeuLeuLeuGlnGln
2041 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG
     GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln
2101 CAGCAGCAAGAGACTAGCCCCAGGCAGCAGCAGCAGCAGGGTGAGGATGGTTCTCCC
     GlnGlnGlnGluThrSerProArgGlnGlnGlnGlnGlnGlyGluAspGlySerPro
2161 CAAGCCCATCGTAGAGGCCCCACAGGCTACCTGGTCCTGGATGAGGAACAGCAACCTTCA
     GlnAlaHisArgArgGlyProThrGlyTyrLeuValLeuAspGluGluGlnGlnProSer
2221 CAGCCGCAGTCGGCCCTGGAGTGCCACCCCGAGAGAGGTTGCGTCCCAGAGCCTGGAGCC
     GlnProGlnSerAlaLeuGluCysHisProGluArgGlyCysValProGluProGlyAla
2281 GCCGTGGCCGCCAGCAAGGGGCTGCCGCAGCAGCTGCCAGCACCTCCGGACGAGGATGAC
     AlaValAlaAlaSerLysGlyLeuProGlnGlnLeuProAlaProProAspGluAspAsp
2341 TCAGCTGCCCCATCCACGTTGTCCCTGCTGGGCCCCACTTTCCCCGGCTTAAGCAGCTGC
     SerAlaAlaProSerThrLeuSerLeuLeuGlyProThrPheProGlyLeuSerSerCys
2401 TCCGCTGACCTTAAAGACATCCTGAGCGAGGCCAGCACCATGCAACTCCTTCAGCAACAG
     SerAlaAspLeuLysAspIleLeuSerGluAlaSerThrMetGlnLeuLeuGlnGlnGln
2461 CAGCAGGAAGCAGTATCCGAAGGCAGCAGCAGCGGGAGAGCGAGGGAGGCCTCGGGGGCT
     GlnGlnGluAlaValSerGluGlySerSerSerGlyArgAlaArgGluAlaSerGlyAla
2521 CCCACTTCCTCCAAGGACAATTACTTAGGGGGCACTTCGACCATTTCTGACAACGCCAAG
     ProThrSerSerLysAspAsnTyrLeuGlyGlyThrSerThrIleSerAspAsnAlaLys
```

FIG. 5B

```
2581 GAGTTGTGTAAGGCAGTGTCGGTGTCCATGGGCCTGGGTGTGGAGGCGTTGGAGCATCTG
     GluLeuCysLysAlaValSerValSerMetGlyLeuGlyValGluAlaLeuGluHisLeu

2641 AGTCCAGGGGAACAGCTTCGGGGGGATTGCATGTACGCCCCACTTTTGGGAGTTCCACCC
     SerProGlyGluGlnLeuArgGlyAspCysMetTyrAlaProLeuLeuGlyValProPro

2701 GCTGTGCGTCCCACTCCTTGTGCCCCATTGGCCGAATGCAAAGGTTCTCTGCTAGACGAC
     AlaValArgProThrProCysAlaProLeuAlaGluCysLysGlySerLeuLeuAspAsp

2761 AGCGCAGGCAAGAGCACTGAAGATACTGCTGAGTATTCCCCTTTCAAGGGAGGTTACACC
     SerAlaGlyLysSerThrGluAspThrAlaGluTyrSerProPheLysGlyGlyTyrThr

2821 AAAGGGCTAGAAGGCGAGAGCCTAGGCTGCTCTGGCAGCGCTGCAGCAGGGAGCTCCGGG
     LysGlyLeuGluGlyGluSerLeuGlyCysSerGlySerAlaAlaAlaGlySerSerGly

2881 ACACTTGAACTGCCGTCTACCCTGTCTCTCTACAAGTCCGGAGCACTGGACGAGGCAGCT
     ThrLeuGluLeuProSerThrLeuSerLeuTyrLysSerGlyAlaLeuAspGluAlaAla

2941 GCGTACCAGAGTCGCGACTACTACAACTTTCCACTGGCTCTGGCCGGACCGCCGCCCCCT
     AlaTyrGlnSerArgAspTyrTyrAsnPheProLeuAlaLeuAlaGlyProProProPro

3001 CCGCCGCCTCCCCATCCCCACGCTCGCATCAAGCTGGAGAACCCGCTGGACTACGGCAGC
     ProProProProHisProHisAlaArgIleLysLeuGluAsnProLeuAspTyrGlySer

3061 GCCTGGGCGGCTGCGGCGGCGCAGTGCCGCTATGGGGACCTGGCGAGCCTGCATGGCGCG
     AlaTrpAlaAlaAlaAlaAlaGlnCysArgTyrGlyAspLeuAlaSerLeuHisGlyAla

3121 GGTGCAGCGGGACCCGGTTCTGGGTCACCCTCAGCCGCCGCTTCCTCATCCTGGCACACT
     GlyAlaAlaGlyProGlySerGlySerProSerAlaAlaAlaSerSerSerTrpHisThr

3181 CTCTTCACAGCCGAAGAAGGCCAGTTGTATGGACCGTGTGGTGGTGGTGGGGGTGGTGGC
     LeuPheThrAlaGluGluGlyGlnLeuTyrGlyProCysGlyGlyGlyGlyGlyGlyGly

3241 GGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGAGGCGGGA
     GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGluAlaGly

3301 GCTGTAGCCCCCTACGGCTACACTCGGCCCCCTCAGGGGCTGGCGGGCCAGGAAAGCGAC
     AlaValAlaProTyrGlyTyrThrArgProProGlnGlyLeuAlaGlyGlnGluSerAsp

3361 TTCACCGCACCTGATGTGTGGTACCCTGGCGGCATGGTGAGCAGAGTGCCCTATCCCAGT
     PheThrAlaProAspValTrpTyrProGlyGlyMetValSerArgValProTyrProSer

3421 CCCACTTGTGTCAAAAGCGAAATGGGCCCCTGGATGGATAGCTACTCCGGACCTTACGGG
     ProThrCysValLysSerGluMetGlyProTrpMetAspSerTyrSerGlyProTyrGly

3481 GACATGCGTTTGGAGACTGCCAGGGACCATGTTTTGCCCATTGACTATTACTTTCCACCC
     AspMetArgLeuGluThrAlaArgAspHisValLeuProIleAspTyrTyrPheProPro
```

FIG. 5C

```
3541 CAGAAGACCTGCCTGATCTGTGGAGATGAAGCTTCTGGGTGTCACTATGGAGCTCTCACA
     GlnLysThrCysLeuIleCysGlyAspGluAlaSerGlyCysHisTyrGlyAlaLeuThr

3601 TGTGGAAGCTGCAAGGTCTTCTTCAAAAGAGCCGCTGAAGGGAAACAGAAGTACCTGTGC
     CysGlySerCysLysValPhePheLysArgAlaAlaGluGlyLysGlnLysTyrLeuCys

3661 GCCAGCAGAAATGATTGCACTATTGATAAATTCCGAAGGAAAAATTGTCCATCTTGTCGT
     AlaSerArgAsnAspCysThrIleAspLysPheArgArgLysAsnCysProSerCysArg

3721 CTTCGGAAATGTTATGAAGCAGGGATGACTCTGGGAGCCCGGAAGCTGAAGAAACTTGGT
     LeuArgLysCysTyrGluAlaGlyMetThrLeuGlyAlaArgLysLeuLysLysLeuGly

3781 AATCTGAAACTACAGGAGGAAGGAGAGGCTTCCAGCACCACCAGCCCCACTGAGGAGACA
     AsnLeuLysLeuGlnGluGluGlyGluAlaSerSerThrThrSerProThrGluGluThr

3841 ACCCAGAAGCTGACAGTGTCACACATTGAAGGCTATGAATGTCAGCCCATCTTTCTGAAT
     ThrGlnLysLeuThrValSerHisIleGluGlyTyrGluCysGlnProIlePheLeuAsn

3901 GTCCTGGAAGCCATTGAGCCAGGTGTAGTGTGTGCTGGACACGACAACAACCAGCCCGAC
     ValLeuGluAlaIleGluProGlyValValCysAlaGlyHisAspAsnAsnGlnProAsp

3961 TCCTTTGCAGCCTTGCTCTCTAGCCTCAATGAACTGGGAGAGAGACAGCTTGTACACGTG
     SerPheAlaAlaLeuLeuSerSerLeuAsnGluLeuGlyGluArgGlnLeuValHisVal

4021 GTCAAGTGGGCCAAGGCCTTGCCTGGCTTCCGCAACTTACACGTGGACGACCAGATGGCT
     ValLysTrpAlaLysAlaLeuProGlyPheArgAsnLeuHisValAspAspGlnMetAla

4081 GTCATTCAGTACTCCTGGATGGGGCTCATGGTGTTTGCCATGGGCTGGCGATCCTTCACC
     ValIleGlnTyrSerTrpMetGlyLeuMetValPheAlaMetGlyTrpArgSerPheThr

4141 AATGTCAACTCCAGGATGCTCTACTTCGCCCCTGATCTGGTTTTCAATGAGTACCGCATG
     AsnValAsnSerArgMetLeuTyrPheAlaProAspLeuValPheAsnGluTyrArgMet

4201 CACAAGTCCCGGATGTACAGCCAGTGTGTCCGAATGAGGCACCTCTCTCAAGAGTTTGGA
     HisLysSerArgMetTyrSerGlnCysValArgMetArgHisLeuSerGlnGluPheGly

4261 TGGCTCCAAATCACCCCCCAGGAATTCCTGTGCATGAAAGCACTGCTACTCTTCAGCATT
     TrpLeuGlnIleThrProGlnGluPheLeuCysMetLysAlaLeuLeuLeuPheSerIle

4321 ATTCCAGTGGATGGGCTGAAAAATCAAAAATTCTTTGATGAACTTCGAATGAACTACATC
     IleProValAspGlyLeuLysAsnGlnLysPhePheAspGluLeuArgMetAsnTyrIle

4381 AAGGAACTCGATCGTATCATTGCATGCAAAAGAAAAAATCCCACATCCTGCTCAAGACGC
     LysGluLeuAspArgIleIleAlaCysLysArgLysAsnProThrSerCysSerArgArg

4441 TTCTACCAGCTCACCAAGCTCCTGGACTCCGTGCAGCCTATTGCGAGAGAGCTGCATCAG
     PheTyrGlnLeuThrLysLeuLeuAspSerValGlnProIleAlaArgGluLeuHisGln
```

FIG. 5D

```
4501 TTCACTTTTGACCTGCTAATCAAGTCACACATGGTGAGCGTGGACTTTCCGGAAATGATG
     PheThrPheAspLeuLeuIleLysSerHisMetValSerValAspPheProGluMetMet

4561 GCAGAGATCATCTCTGTGCAAGTGCCCAAGATCCTTTCTGGGAAAGTCAAGCCCATCTAT
     AlaGluIleIleSerValGlnValProLysIleLeuSerGlyLysValLysProIleTyr

4621 TTCCACACCCAGTGAAGCATTGGAAACCCTATTTCCCCACCCCAGCTCATGCCCCCTTTC
     PheHisThrGlnEnd

4681 AGATGTCTTCTGCCTGTTATAACTCTGCACTACTCCTCTGCAGTGCCTTGGGGAATTTCC

4741 TCTATTGATGTACAGTCTGTCATGAACATGTTCCTGAATTCTATTTGCTGGGCTTTTTTT

4801 TTCTCTTTCTCTCCTTTCTTTTTCTTCTTCCCTCCCTATCTAACCCTCCCATGGCACCTT

4861 CAGACTTTGCTTCCCATTGTGGCTCCTATCTGTGTTTTGAATGGTGTTGTATGCCTTTAA

4921 ATCTGTGATGATCCTCATATGGCCCAGTGTCAAGTTGTGCTTGTTTACAGCACTACTCTG

4981 TGCCAGCCACACAAACGTTTACTTATCTTATGCCACGGGAAGTTTAGAGAGCTAAGATTA

5041 TCTGGGGAAATCAAAACAAAAAACAAGCAAACAAAAAAAAAA 5082
```

FIG. 5E

```
AATTCGGGAAGGATCGAGCAAACCAGGAAAGTAAGGATGGAGATCCTAGGAGAGTGTCCA    60

TGCCTCGAAAGGAGCCCACCAAAGATGAACTGTTGCATTTGCTTTCCACCTCCCAGCGCC   120

CCCTCGGAGATCCCTAGGAGCCAGCCTGCTGGGAGAACCAGAGGGTCCGGAGCAAACCTG   180

GAGGCTGAGAGGGCATCAGAGGGGAAAAGACTGAGTTAGCCACTCCAGTGCCATACAGAA   240

GCTTAAGGGACATACCACGCCAGCCCCAGCCCAGCGACAGCCAACGCCTGTTGCAGAGCG   300

GCGGCTTCGAAGCCGCCGCCCAGAAGCTGCCCTTTCCTCTTCGGTGAAGTTTCTAAAAGC   360

TGCGGGAGACTCGGAGGAAGCGAAGAAAGTGTCCGGTAGGACTACGACTGCCTTTGTCCT   420

CCTCCCTCCTACCCCTACCCCTCCTGGGTCCCCTCTCCCTGAGCGGACTAGGCAGGCTTC   480

CTGGCCAGCCCTCTCCCCTACACCACCAGCTCTGCCAGCCAGTTTGCACAGAGGTAACTC   540

CCTTTGGCTGAAAGCAGACGAGCTTGTTGCCCATTGGAAGGGAGGCTTTTGGGAGCCCAG   600

AGACTGAGGAGCAACAGCACGCTGGAGAGTCCCTGATTCCAGGTTCTCCCCCCTGCACCT   660

CCTACTGCCCGCCCCTCACCCTGTGTGTGCAGCTAGAATTGAAAAGATGAAAAGACAGTT   720

GGGGCTTCAGTAGTCGAAAGCAAAACAAAAGCAAAAAGAAAACAAAAAGAAAATAGCCCA   780

GTTCTTATTTGCACCTGCTTCAGTGGACATTGACTTTGGAAGGCAGAGAATTTTCCTTCC   840

CCCCAGTCAAGCTTTGAGCATCTTTTAATCTGTTCTTCAAGTATTTAGGGACAAACTGTG   900

AAACTAGCAGGGCAGATCCTGTCTAGCGCGTGCCTTCCTTTACAGGAGACTTTGAGGCTA   960

TCTGGGCGCTCCCCCCCCTCCCTGCAAGTTTTCTTCCCTGGAGCTTCCCGCAGGTGGGCA  1020

GCTAGCTGCAGATACTACATCATCAGTCAGTAGAACTCTTCAGAGCAAGAGACGAGGAGG  1080

CAGGATAAGGGAATTCGGTGGAAGCTAGAGACAAGCTAAAGGATGGAGGTGCAGTTAGGG  1140
                                                MetGluValGlnLeuGly

CTGGGAAGGGTCTACCCACGGCCCCCGTCCAAGACCTATCGAGGAGCGTTCCAGAATCTG  1200
LeuGlyArgValTyrProArgProProSerlysThrTyrArgGlyAlaPheGlnAsnLeu

TTCCAGAGCGTGCGCGAAGCGATCCAGAACCCGGGCCCCAGGCACCCTGAGGCCGCTAGC  1260
PheGlnSerValArgGluAlaIleGlnAsnProGlyProArgHisProGluAlaAlaSer

ATAGCACCTCCCGGTGCCTGTTTACAGCAGCGGCAGGAGACTAGCCCCCGGCGGCGGCGG  1320
IleAlaProProGlyAlaCysLeuGlnGlnArgGlnGluThrSerProArgArgArgArg

CGGCAGCAGCACCCTGAGGATGGCTCTCCTCAAGCCCACATCAGAGGCACCACAGGCTAC  1380
ArgGlnGlnHisProGluAspGlySerProGlnAlaHisIleArgGlyThrThrGlyTyr
```

FIG. 6A

```
CTGGCCCTGGAGGAGGAACAGCAGCCTTCACAGCAGCAGTCAGCCTCCGAGGGCCACCCT 1440
LeuAlaLeuGluGluGluGlnGlnProSerGlnGlnGlnSerAlaSerGluGlyHisPro

GAGAGCGGCTGCCTCCCGGAGCCTGGAGCTGCCACGGCTCCTGGCAAGGGGCTGCCGCAG 1500
GluSerGlyCysLeuProGluProGlyAlaAlaThrAlaProGlyLysGlyLeuProGln

CAGCCACCAGCTCCTCCAGATCAGGATGACTCAGCTGCCCCATCCACGTTGTCCCTACTG 1560
GlnProProAlaProProAspGlnAspAspSerAlaAlaProSerThrLeuSerLeuLeu

GGCCCCACTTTCCCAGGCTTAAGCAGCTGCTCCGCAGACATTAAAGACATCCTGAGCGAG 1620
GlyProThrPheProGlyLeuSerSerCysSerAlaAspIleLysAspIleLeuSerGlu

GCCGGCACCATGCAACTTCTTCAGCAGCAGCAGCAACAGCAACAGCAGCAGCAGCAGCAG 1680
AlaGlyThrMetGlnLeuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln

CAGCAGCAGCAGCAGCAACAGCAGCAGGAGGTAATATCCGAAGGCAGCAGCAGCGTGAGA 1740
GlnGlnGlnGlnGlnGlnGlnGlnGlnGluValIleSerGluGlySerSerSerValArg

GCAAGGGAGGCCACTGGGGCTCCCTCTTCCTCCAAGGATAGTTACCTAGGGGGCAATTCG 1800
AlaArgGluAlaThrGlyAlaProSerSerSerLysAspSerTyrLeuGlyGlyAsnSer

ACCATATCTGACAGTGCCAAGGAGTTGTGTAAAGCAGTGTCTGTGTCCATGGGGTTGGGT 1860
ThrIleSerAspSerAlaLysGluLeuCysLysAlaValSerValSerMetGlyLeuGly

GTGGAAGCACTGGAACATCTGAGTCCAGGGGAGCAGCTTCGGGGCGACTGCATGTACGCG 1920
ValGluAlaLeuGluHisLeuSerProGlyGluGlnLeuArgGlyAspCysMetTyrAla

TCGCTCCTGGGAGGTCCACCCGCCGTGCGTCCCACTCCTTGTGCGCCTCTGGCCGAATGC 1980
SerLeuLeuGlyGlyProProAlaValArgProThrProCysAlaProLeuAlaGluCys

AAAGGTCTTTCCCTGGACGAAGGCCCGGGCAAAGGCACTGAAGAGACTGCTGAGTATTCC 2040
LysGlyLeuSerLeuAspGluGlyProGlyLysGlyThrGluGluThrAlaGluTyrSer

TCTTTCAAGGGAGGTTACGCCAAAGGGTTGGAAGGTGAGAGTCTGGGCTGCTCTGGCAGC 2100
SerPheLysGlyGlyTyrAlaLysGlyLeuGluGlyGluSerLeuGlyCysSerGlySer

AGTGAAGCAGGTAGCTCTGGGACACTTGAGATCCCGTCCTCACTGTCTCTGTATAAGTCT 2160
SerGluAlaGlySerSerGlyThrLeuGluIleProSerSerLeuSerLeuTyrLysSer

GGAGCAGTAGACGAGGCAGCAGCATACCAGAATCGCGACTACTACAACTTTCCGCTCGCT 2220
GlyAlaValAspGluAlaAlaAlaTyrGlnAsnArgAspTyrTyrAsnPheProLeuAla

CTGTCCGGGCCGCCGCACCCCCCGCCCCCTACCCATCCACACGCCCGCATCAAGCTGGAG 2280
LeuSerGlyProProHisProProProProThrHisProHisAlaArgIleLysLeuGlu

AACCCGTCGGACTACGGCAGCGCCTGGGCTGCGGCGGCAGCGCAATGCCGCTATGGGGAC 2340
AsnProSerAspTyrGlySerAlaTrpAlaAlaAlaAlaAlaGlnCysArgTyrGlyAsp

TTGGCTAGCCTACATGGAGGGAGTGTAGCCGGACCCAGCACTGGATCGCCCCCAGCCACC 2400
LeuAlaSerLeuHisGlyGlySerValAlaGlyProSerThrGlySerProProAlaThr
```

FIG. 6B

GCCTCTTCTTCCTGGCATACTCTCTTCACAGCTGAAGAAGGCCAATTATATGGGCCAGGA 2460
AlaSerSerSerTrpHisThrLeuPheThrAlaGluGluGlyGlnLeuTyrGlyProGly

GGCGGGGGCGGCAGCAGTAGCCCAAGCGATGCTGGGCCTGTAGCCCCCTATGGCTACACT 2520
GlyGlyGlyGlySerSerSerProSerAspAlaGlyProValAlaProTyrGlyTyrThr

CGGCCCCCTCAGGGGCTGGCAAGCCAGGAGGGTGACTTCTCTGCCTCTGAAGTGTGGTAT 2580
ArgProProGlnGlyLeuAlaSerGlnGluGlyAspPheSerAlaSerGluValTrpTyr

CCTGGTGGAGTTGTGAACAGAGTCCCCTATCCCAGTCCCAGTTGTGTTAAAAGTGAAATG 2640
ProGlyGlyValValAsnArgValProTyrProSerProSerCysValLysSerGluMet

GGACCTTGGATGGAGAACTACTCCGGACCTTATGGGGACATGCGTTTGGACAGTACCAGG 2700
GlyProTrpMetGluAsnTyrSerGlyProTyrGlyAspMetArgLeuAspSerThrArg

GACCACGTTTTACCCATCGACTATTACTTCCCACCCCAGAAGACCTGCCTGATCTGTGGA 2760
AspHisValLeuProIleAspTyrTyrPheProProGlnLysThrCysLeuIleCysGly

GATGAAGCTTCTGGTTGTCACTACGGAGCTCTCACTTGTGGCAGCTGCAAGGTCTTCTTC 2820
AspGluAlaSerGlyCysHisTyrGlyAlaLeuThrCysGlySerCysLysValPhePhe

AAAAGAGCTGCGGAAGGGAAACAGAAGTATCTATGTGCCAGCAGAAATGATTGCACCATT 2880
LysArgAlaAlaGluGlyLysGlnLysTyrLeuCysAlaSerArgAsnAspCysThrIle

GATAAATTTCGGAGGAAAAATTGTCCATCGTGTCGTCTCCGGAAATGTTATGAAGCAGGG 2940
AspLysPheArgArgLysAsnCysProSerCysArgLeuArgLysCysTyrGluAlaGly

ATGACTCTGGGAGCTCGTAAGCTGAAGAAACTTGGAAATCTCAAACTACAGGAAGAAGGA 3000
MetThrLeuGlyAlaArgLysLeuLysLysLeuGlyAsnLeuLysLeuGlnGluGluGly

GAAAACTCCAGTGCTGGTAGCCCCACTGAGGACCCATCCCAGAAGATGACTGTATCACAC 3060
GluAsnSerSerAlaGlySerProThrGluAspProSerGlnLysMetThrValSerHis

ATTGAAGGCTATGAATGTCAACCTATCTTTCTTAATGTCCTGGAAGCCATTGAGCCAGGA 3120
IleGluGlyTyrGluCysGlnProIlePheLeuAsnValLeuGluAlaIleGluProGly

GTGGTGTGTGCCGGACATGACAACAACCAGCCTGATTCCTTTGCTGCCTTGTTATCTAGT 3180
ValValCysAlaGlyHisAspAsnAsnGlnProAspSerPheAlaAlaLeuLeuSerSer

CTCAACGAGCTTGGCGAGAGACAGCTTGTACATGTGGTCAAGTGGGCCAAGGCCTTGCCT 3240
LeuAsnGluLeuGlyGluArgGlnLeuValHisValValLysTrpAlaLysAlaLeuPro

GGCTTCCGCAACTTGCATGTGGATGACCAGATGGCAGTCATTCAGTATTCCTGGATGGGA 3300
GlyPheArgAsnLeuHisValAspAspGlnMetAlaValIleGlnTyrSerTrpMetGly

CTGATGGTATTTGCCATGGGTTGGCGGTCCTTCACTAATGTCAACTCTAGGATGCTCTAC 3360
LeuMetValPheAlaMetGlyTrpArgSerPheThrAsnValAsnSerArgMetLeuTyr

TTTGCACCTGACCTGGTTTTCAATGAGTATCGCATGCACAAGTCTCGAATGTACAGCCAG 3420
PheAlaProAspLeuValPheAsnGluTyrArgMetHisLysSerArgMetTyrSerGln

FIG. 6C

```
TGCGTGAGGATGAGGCACCTTTCTCAAGAGTTTGGATGGCTCCAGATAACCCCCCAGGAA 3480
CysValArgMetArgHisLeuSerGlnGluPheGlyTrpLeuGlnIleThrProGlnGlu

TTCCTGTGCATGAAAGCACTGCTACTCTTCAGCATTATTCCAGTGGATGGGCTGAAAAAT 3540
PheLeuCysMetLysAlaLeuLeuPheSerIleIleProValAspGlyLeuLysAsn

CAAAAATTCTTTGATGAACTTCGAATGAACTACATCAAGGAACTTGATCGCATCATTGCA 3600
GlnLysPhePheAspGluLeuArgMetAsnTyrIleLysGluLeuAspArgIleIleAla

TGCAAAAGAAAAAATCCCACATCCTGCTCAAGGCGCTTCTACCAGCTCACCAAGCTCCTG 3660
CysLysArgLysAsnProThrSerCysSerArgArgPheTyrGlnLeuThrLysLeuLeu

GATTCTGTGCAGCCTATTGCAAGAGAGCTGCATCAATTCACTTTTGACCTGCTAATCAAG 3720
AspSerValGlnProIleAlaArgGluLeuHisGlnPheThrPheAspLeuLeuIleLys

TCCCATATGGTGAGCGTGGACTTTCCTGAAATGATGGCAGAGATCATCTCTGTGCAAGTG 3780
SerHisMetValSerValAspPheProGluMetMetAlaGluIleIleSerValGlnVal

CCCAAGATCCTTTCTGGGAAAGTCAGCCCATGTATTTCCACACACAGTGAAGATTTGGAA 3840
ProLysIleLeuSerGlyLysValSerProCysIleSerThrHisSerGluAspLeuGlu

CCTAATACCCAAACCCACCTGTTCCCTTTTCAGATGTCTTCTGCCTGTTATATAACTCTG 3900
ProAsnThrGlnThrHisLeuPheProPheGlnMetSerSerAlaCysTyrIleThrLeu

CACTACTTCTCTGGCATGGGCCTTGGGGGAAATTCCTCTACTGATGTACAGTCTGTCATG 3960
HisTyrPheSerGlyMetGlyLeuGlyGlyAsnSerSerThrAspValGlnSerValMet

AACATGTTCCCCAAGTTCTATTTCCTGGGCTTTTCCTTCTTTCTTTTTCTTCTTCTCTGC 4020
AsnMetPheProLysPheTyrPheLeuGlyPheSerPhePheLeuPheLeuLeuLeuCys

CTCTTTTACCCTCCCATGGCACATTTTGAATCCGCTGCGTGTTGTGGCTCCTGCCTGTGT 4080
LeuPheTyrProProMetAlaHisPheGluSerAlaAlaCysCysGlySerCysLeuCys

TTTGAGTTTTGTTGTATTTCTTCAAGTCTGTGATGATCTTCTTGTGGCCCAGTGTCAACT 4140
PheGluPheCysCysIleSerSerSerLeuEnd

GTGCTTGTTTATAGCACTGTGCTGTGTGCCAACCAAGCAAATGTTTACTCACCTTATGCC 4200

ATGGCAAGTTTAGAGAGCTATAAGTATCTTGGGAAGAAACAAACAGAGAGAGTAAAAAAA 4260

CCAAAAAAAAAAAAAAAAAACCGAATTC                                  4288
```

FIG. 6D

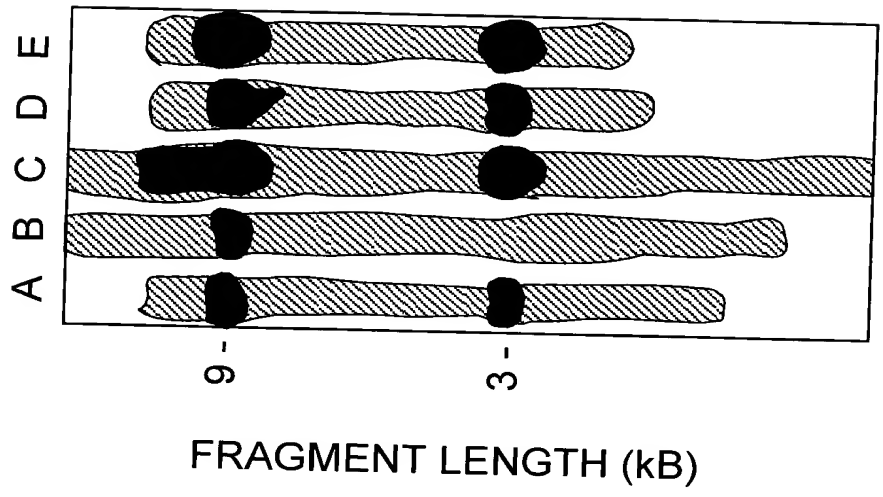

A B C D E
9 -
3 -
FRAGMENT LENGTH (kB)

FIG. 8

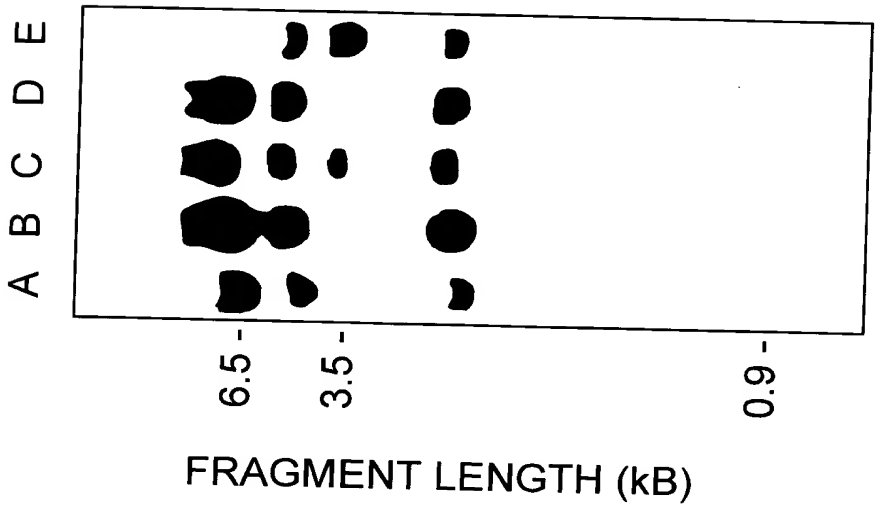

A B C D E
6.5 -
3.5 -
0.9 -
FRAGMENT LENGTH (kB)

FIG. 7